Published in final edited form as:
J Natl Med Assoc. 2008 October ; 100(10): 1235–1243.

Conceptualizing Race in Research

Giselle Corbie-Smith, MD, MSc, **Gail Henderson, PhD**, **Connie Blumenthal, MPH**, **Jessica Dorrance**, and **Sue Estroff, PhD**
Department of Social Medicine (Corbie-Smith, associate professor; Henderson, professor; Dorrance, research assistant; Estroff, professor), Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill; Program on Health Disparities, Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill (Corbie-Smith, associate professor; Blumenthal, research associate; Dorrance, research assistant), Chapel Hill, NC

Abstract

Background—The use of race as a variable in research continues to spark debate about whether it should be used, as well as the implications it has for research on health differences. Given this continued controversy, we examined how investigators interpret the concept of "race" and whether their views of race are reflected in their published work.

Methods—Thirty-three semistructured interviews were conducted with investigators from 3 southeastern universities to discuss recruitment of participants, the use of race as a variable in research and analyses, and their assessment of the National Institutes of Health mandate on the inclusion of women and minorities. The interview data were analyzed using the principles of constant comparative method, theme identification and pattern investigation. Up to 2 publications for each respondent were also used to assess the use of race in their research.

Results—Results reflect a spectrum of views on the definition of race, from biological to social. Findings also suggest that investigators think critically about the use and implications of using race in their research, although this is not consistently reflected in their published work.

Conclusion—In our view, authors, journal editors and peer reviewers have an important role in moving this debate forward, and advocate that they engage more directly in shaping the process. When reporting results by race, investigators should provide a statement on the theory or conceptual framework underlying the hypothesized racial differences in health examined in the study. They should be also cautious in invoking either biological or social constructions of race, thus demonstrating an appreciation of the nuances and implications of using this variable.

Keywords

race/ethnicity; minority health; health disparities

Research on disparities in health is in transition from largely descriptive investigations of the existence of racial differences to analysis of potential etiologies of these variations.[1,2] However, identifying the causal pathways that link race to health outcomes has been complicated by debate over whether race should be used as a variable in such research, and if it is used, whether it is a biological construct, a social construct or something in between.[3–13]

Send correspondence and reprint requests to: Dr. Giselle Corbie-Smith, Department of Social Medicine, University of North Carolina at Chapel Hill, Medical School, Wing D, CB#7240, Chapel Hill, NC 27599; phone: (919) 962-1136; (919) 843-6877; fax: (919) 966-7499; gcorbie@med.unc.edu.

These debates have been fueled by the need for researchers to respond to mandated inclusion of racial and ethnic minorities in federally funded research, changes in Office of Management and Budget (OMB) categories of race and ethnicity, and closer examination of the use of race as a variable by editors of scientific journals. The National Institutes of Health (NIH) policy on the inclusion of women and minorities in clinical research requires investigators funded by the federal government to document the participation of these groups or to make a sound scientific argument for their exclusion.[14,15] Similarly, federally funded research must comply with OMB Directive 15 revised standards on racial and ethnic classification (*Federal Register*, 1997), which separate race from ethnicity (specifically, Hispanic origin). These requirements for reporting racial and ethnic categories apply to research with human subjects whether or not an investigator is focusing on racial or ethnic differences in health.

One consequence of these co-occurring events is that it has become easier to make race-based comparisons. The scientific inquiry into and resultant literature examining associations of race and health outcomes have increased significantly in recent years.[2] However, the authors of this literature seldom address why "race" is a key independent variable for analysis, and if it is being used as a proxy for social class, cultural background, biological difference or some combination.[3,4,16,17] These concerns have prompted the editorial boards of scientific journals to question the common practice of reporting analyses by race without further justification (e.g., *Journal of the American Medical Association*, *Nature Genetics*, *Archives of Internal Medicine*, *American Psychologist* and *British Medical Journal*).

At present, investigators who examine and report racial differences in health must make decisions about the relevance and application of race in their work, raising questions about how data on race should be collected, how federal requirements should guide recruitment strategies, and what derivative implications for research design and analysis arise. Given continued controversy surrounding the use of race as a variable and the implications for research on health differences by race, we examined how investigators interpret the concept of "race" and whether their views of race are reflected in their published work.

METHODS

Subject Identification and Recruitment

We sampled investigators conducting funded clinical research involving adult human subjects in the year 2000 from 3 southeastern universities: 1 historically black university (HBCU) and 2 research intensive universities (1 private and 1 public). We identified investigators by using institutional review board lists of all active studies during that year, from each institution, searching the university websites and the Computer Retrieval of Information on Scientific Projects (CRISP) database, a searchable database of NIH-funded biomedical research projects, maintained by the NIH Office of Extramural Research. Using the title of the study and publicly available descriptions (e.g., abstracts), the principal investigator (PI) list was stratified by institution, whether the investigator had a history of NIH funding and whether their work focused on minority health or race-related health disparities. The investigator sample was generated using random selection within these 12 strata. We considered institutional culture, NIH-mandated inclusion of minorities and the personal choice of a career focus on the health of minority populations as well as these variables, as potentially directly or indirectly impacting an investigators' concepts of race. Because the number of investigators at the HBCU conducting clinical research initially identified was small, we also used snowball sampling at that site (i.e., we asked each respondent to identify other potentially study-eligible investigators at their institution) and attempted to interview all investigators at that university who met our inclusion criteria.

Investigators were first contacted by a letter containing a brief description of the project and an invitation to participate in the interview. Using a prepared recruitment script, investigators were then contacted by telephone to further explain the nature of the study and verify eligibility. Up to 4 telephone and e-mail attempts were made to contact investigators. After no response to 4 contact attempts, the respondent was considered a refusal and replaced with another investigator randomly chosen from the sample for that institution.

Data Collection

Investigator interviews—Data were collected by in-depth telephone interviews. Using a semistructured guide, the interviewer asked open-ended questions, followed up participants' responses, pursued themes as they arose, and sought clarification or elaboration as needed. Interviews were conducted between November 2001 and January 2003. All interviews were conducted by the first author (GCS) and were 30–50 minutes in length. The interview guide covered the following areas: the participant's research topics, recent experience with participant recruitment, influences on recruitment of minority subjects, use of race as a variable in research and analyses, and assessing any influence of the NIH mandate on the conduct of the research. In the interview guide for this study, we specifically and consistently used the term "race" rather than "race/ethnicity" or "ethnicity." Since the latter terms have been heavily debated and endorsed by some in the scientific literature, we avoided introducing a concept or idea that might lead to socially desirable responses from the interview participants. We did not ask investigators whether they saw race as a biologic or social construct. Interviews were audiotaped and transcribed for analysis. Two members of the research team reviewed each interview transcript to verify the accuracy of the transcription and to remove all identifiers. If necessary, a third listener, the interviewer, reviewed the transcript to resolve conflicts or areas of confusion in the transcription.

Investigator publications—Publication of research represents the basic data for the scientific and conceptual life of race, and the impact of recent journal guidelines regarding race as a variable was of interest. As a result, we examined up to 2 publications for each respondent, limiting the sample to those that were retrievable through Pub Med/MedLine. Analysis of their use of race in these publications was also a means to both expand upon and triangulate interview findings. The goal was to match ≥1 published article with the study described by the PI in the interview. If there were no published articles matching the study description provided by the PI, we searched for an article that included the following MeSH terms or keywords: racial stocks (for example: "race," "African American," "black," etc). In addition, when possible, ≥1 article where the PI was the first, last or second author was included. Only articles published since 1998 were included and when >1 article met all the inclusion criteria, the most recent article was selected. No published letters or commentaries were included in this analysis, but review articles were included. We used a data abstraction form to record the following information: whether race was mentioned at any point in the manuscript and where (title, introduction, methods, tables or results, discussion); whether race was defined or explained in the manuscript and whether that description fit a social construct and/or biological construct or could not be determined. We also noted the journal in which the paper was published and whether that journal had published guidelines on reporting race (either in the instructions for authors or published statement) at the time of this study.

Analysis

Investigator interviews—Interview transcripts were analyzed using the principles of grounded theory[18] and the content analysis techniques of theme identification. [19,20] Glaser's method of constant comparative analysis requires the data to be reviewed in light of an initial conceptual formulation and coded repeatedly. Codes about the meaning and

measurement of race were initially based on distinctions made in the scientific literature (i.e., social or biological constructs) and expanded inductively using an iterative process. At regular meetings, the research team reviewed all new transcripts for emerging themes, extant codes were revisited, and the coding scheme refined. Written definitions for each code were developed and revised based on input from all research team members. Examples and directions of when and when not to use the codes were detailed in the code book. Sample size was not formally calculated. Instead, participants were enrolled until no new concepts arose during analysis of successive interviews, a concept called theoretical saturation.

After finalizing the coding strategy, 3 research team members (GCS, CB and JD) then coded all transcripts in teams of 2 coders. In addition, 2 senior members of the research team (GEH and SEE) coded randomly selected transcripts the codes reconciled. In this form of analytical triangulation, important insights can emerge from the different ways different analysts look at the same set of data and also provide a means to ensure internal validity.[21] We used ATLAS.ti software to facilitate organization, management and analysis of the qualitative dataset. Coded items were then compared within and between interviews. Participants' comments were sorted by content areas. We selected quotes illustrative of each domain and those that were exemplars of various perspectives within that area.

Investigator publications—Publications of investigator–subjects were coded by 2 coders. The research team deductively developed a set of codes derived from content areas noted in the interviews. The code book included definitions, examples and rules of application. The publications were coded for whether race or racial categories were mentioned in the title, introduction, methods, results, tables or discussion section. If mentioned, the passage was coded, documenting whether a definition or explanation was provided. We compared investigator comments on race in their interviews with their use of race in their published work. Manuscript coders were blind to the codes assigned during the analysis of the investigator interviews. We also examined the instructions for authors for the journals from which publications were selected.

RESULTS

We contacted 43 investigators and conducted 33 interviews (response rate=77%) who had been a PI on between 2–50 (range) prior studies. Characteristics of the investigator–participants in this study are presented in Table 1. When discussing race in their research, the investigators responded in a variety of ways to the questions: "How do you think about race in your research? How do you conceptualize race in your work?" Some investigators focused on issues of the measurement and race in their work, while others went on to describe how they were using the variable, what they thought it meant or how they would explain the finding of racial differences in their work. Each of these themes is described below, and illustrative quotes are shown in Table 2.

Measurement and Race

Investigators described a variety of methods for collecting data on racial categories. Most investigators used self-identification (i.e., asking the research subject) or self-selection based on predefined categories as the main data collection strategy for race. Some investigators augmented these methods in 2 ways: 1) gathering data on self-reported racial identity or using a validated measure of racial identity; or 2) "quantifying" race by asking for detailed information about the race of parents and grandparents, then using a predetermined algorithm to determine the subjects' race. Most investigators also described data collected on race from administrative databases as being unreliable, being uncertain about how to

analyze data from individuals who self-identify as “mixed race” or “other,” and not being able to adequately represent heterogeneity within racial groups.

Race as a Biological Construct

In discussing how race might operate in their research, several investigators endorsed the idea of race as a biologic construct. Responses in this category included the identification of genetic variants, differences in response to treatments and race as a risk marker for disease. Several investigators cited scientific literature to support their hypotheses about why there might be biologic differences between races. For example, they referred to findings on differences in prevalence of hypertension and other chronic diseases or differences in allele frequencies.

Race as a Social Construct

Many investigators described their conceptualization of race as a social construct rather than biologic or genetic. Most investigators in this category pointed to differences in risk of disease based on social factors that are associated with race (e.g., environmental, behavior, diet, socioeconomic status, education and/or wealth). Some investigators linked race to discrimination, oppression and racism, or invoked feminist theory in their conceptualization of race. One investigator who described race as a multidimensional concept included measures of “ethnic identity” as well as race and detailed data on social position, wealth, ethnic identity and acculturation in the analyses.

Race as Both Social and Biological

Several investigators held both social and biological concepts of race, not as mutually exclusive but as complementary. Those who took this stance varied along a continuum of how much influence they attributed to biological and social components of their model. Here too, investigators referred to extant literature to support their concepts of race. Their frameworks included balanced combinations of social and biological elements, and at times, models that privileged the importance of one concept over the other: for example, some held a primarily biologic model of race and racial differences in health augmented by social factors such as behavior and environment. Others acknowledged possible genetic differences, but said social factors were the primary determinants of racial differences. A third group of respondents suggested both social and biologic factors were equally responsible for racial differences in health and disease.

We also included in this category investigators who used conceptually different ways of collecting data on race and conceptualizing race in their work—for example, using self-identity in data collection and examining genetic differences in drug metabolism.

There were no clear associations between investigator conceptualizations of race and NIH funding, institution, investigator self-identified race or research focus when examined bivariately.

Challenges and Opportunities

Investigators often expressed confusion and/or frustration as they described how they conceptualized race in their work. They noted the challenge of disaggregating race and socioeconomic status, and many endorsed the view of one investigator who observed, that race was “a useless variable, encompassing so many things.” Several investigators also mentioned the challenge of finding and using appropriate measures and scales that had acceptable psychometric and biometric properties, and had been adapted for different racial and ethnic groups. Others raised concerns about the possibilities of misuse of data that

examine racial differences in health through either "blaming the victim" or perpetuating notions of genetic inferiority.

In addition to the challenges described above, many study participants identified opportunities that arose from examining racial differences. They pointed to institutional support for race-focused work, capitalizing on the renewed interest of funding agencies in this area of research, and the possibility of uncovering sources of racial differences in outcomes. Those who used race as "a constant rather than a variable in their research" stated that research that focused on 1 racial group would more readily lead to improved health for a particular population.

Investigators' Use of "Race" in Manuscripts

We identified 50 manuscripts that fit the inclusion criteria. No articles were found for 3 PIs and a single article was found for one PI (Table 3). As noted in Table 3, 84% of the manuscripts mentioned race in the text. Among those, only 18% provided a definition or rationale for how or why race was included as a variable in the study. At the time of this study, none of the journals in the sample gave explicit instructions to authors on how to use race in submitted articles. However, 10% of the journals referred to the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, which asks for description of the sample, suggests avoiding the term "race" and asks investigators to "specify what the descriptors mean."

In reviewing the descriptions and explanations of the use of race as a variable in these manuscripts, we found examples of the 3 categories described above: race as a social construct, biological construct and combined. Two-thirds of the manuscripts in which a definition or explanation were in evidence were coded as consistent with the definition that was coded in the interview transcript.

DISCUSSION

In this study, the investigators we interviewed were thinking critically about the use and implications of using race in their research but did not consistently include this reflection in their published work. Our findings reflect a spectrum of views on the concept of race in health-related research and a range of approaches for collecting data on race. The study participants' conceptualization of race included constructs across a continuum of definitions, from biological to social. While the majority did not define race in their publications, most of those few who *did* used definitions that were consistent with how they described race in the interviews.

These results must be interpreted in light of the limitations of this study. While we sought to sample investigators from institutions with different educational and research missions, and investigators with a variety of funding and research foci, all of the institutions were located in the southeastern United States. Based on the historical and social context of this area, it is possible that investigators in this geographic region may have given more thought to the operationalization of race. However, it is difficult to assess the direction of the potential bias, since it is equally likely that they may have been more or less likely to share their thoughts on race with the interviewer, an African-American female clinician researcher whose work focuses on minority populations. Because of the potential bias that the interviewer's characteristics may have introduced, we conducted all interviews by phone unless an in-person interview was explicitly requested by the respondent. Also, while we used a comprehensive algorithm to select manuscripts, we did not ask investigators to identify articles that reflected their concept of race; the latter strategy may have led to the

inclusion of articles with more explicit descriptions of race or to more concordance between interviews and manuscripts.

Despite these limitations, our results reflect the evolving debate on where and how to use race in health-related research. Authors in the public health, social science and medical literature have described the increasing but also contested use of "race" and "ethnicity" in health research, and are engaged in the debate on how race should be used to understand and explain health disparities.[3–7,9,11–13,22–24] This ongoing discussion in the published literature was well known to our respondents, many of whom cited competing literature to bolster the conceptualization of race in their work, regardless of where they positioned themselves on the biological-social continuum.

The dominant and enduring understanding of race in the biomedical literature is of physiological and biological differences between population groups, particularly those of skin color and facial features. While most researchers no longer subscribe to simplistic ideas of racial differences promoted by eugenicists, an assumption that racial categories are based in part on natural and genetic distinctions underlies much medical research.[4,13,25] This interpretation has been reinforced and supported by advances in genetic research. Investigators call on work in population genetics and use ancestral tree diagrams to support racial classification schemes.[26–31] In addition, genetic variation research has uncovered millions of genetic polymorphisms, [32–34] and a growing body of research seeks to use these variations to make sense of racial differences in health. For example, pharmacogenomics and targeted drug therapy are growing areas of research and are held out as one way to address health disparities.[24,35,36] This is still a controversial position, however, as shown in the arguments over potential racial differences in the effectiveness of heart failure medications[37,38] and birthweight differences between black and white babies.[39–42]

Race is also conceptualized as a social category, emphasizing shared social and cultural heritage and high-lighting the deleterious impact of power differentials that exist in society. [8,12,43] The social interpretation emphasizes the importance of history, geographic origins, language, cultural norms and practices, and religious traditions of a group of people based on common ancestry. Because of this emphasis, some researchers prefer to use the term "ethnicity" instead of "race."[44–46] Proponents of the social construction of race also point to the fact that the idea of human races predated modern genetics and was used to create a hierarchy reflecting notions of moral and social superiority and inferiority. This perspective views race as proxy for the effects of racism, differences in socio-economic status, exposure to the health risks of highly stressful home and neighborhood contexts, less access to care and other life experiences that provide limited opportunities for well-being.[5,7,12,47]

Other authors have noted the variability in how investigators use race to describe the study samples in their research[16,17,48] and found that study samples are often incompletely described. If described, little rationale is given for how racial categories are applied. This variation in the use of race in the scientific literature has led authors to question whether investigators are thinking critically about this variable.[17] In contrast, the investigators we studied struggle with the concept of race in their research, describing a range of methods to collect data on this variable and considering the "best" application in their work. Investigators were able to articulate some of the methodological challenges and opportunities in using race as a variable in their analyses.

In our view, authors, journal editors and peer reviewers have an important role in moving this debate forward, and we recommend that they play a more explicit and widespread role in shaping the process. Investigators in our study referred to the scientific literature as they contemplated the meaning of race in their work. If peer reviewers and journal editors

demand that authors clarify and refine definitions and theories, the dialogue will be advanced, as will continued reporting of empirical findings that explicitly define causal pathways linking race and racial differences in health.

When reporting results by race, investigators should provide a statement on the theory underlying hypothesized racial differences in health. They should be cautious in invoking either biological or social constructions of race, thus demonstrating an appreciation of the nuances and implications of using this variable. In addition, as is evident from our findings, investigators hold a range of views on the concept of race, and authors should not assume that others know what is meant by the term "race" in their published work. Journals should require authors to provide explicit definitions of this term, following the lead of several journals. (*Journal of the American Medical Association*, *Nature Genetics*, *Archives of Internal Medicine*, *American Psychologist*, *British Medical Journal*). The researchers we interviewed were grappling with the current unsettled status of race and ethnicity, both empirically and conceptually. Yet these observations, reservations and critiques remain largely "backstage" or informal among colleagues. In order to make progress regarding race as a factor in health, we suggest that a critical and reflective analysis of how race was conceptualized and measured be included as an important component of discussion sections of publications. In conclusion, as we continue to work to understand and address differences in health by race, attention to the complex measurement issues and clear descriptions of underlying theory of how race is conceptualized in a body of research will advance the debate in this field.

Acknowledgments

Financial support: This project was supported by grants from the National Heart Lung and Blood Institute, the Robert Wood Johnson Foundation Harold Amos Medical Faculty Development Program, the National Center for Minority Health and Health Disparities, and an unrestricted grant from Pfizer Inc.

References

1. LaVeist TA. On the study of race, racism, and health: A shift from description to explanation. Int J Health Serv 2000;30(1):217–219. [PubMed: 10707307]
2. Smedley, BD.; Stith, AY.; Nelson, AR., editors. Unequal treatment: Confronting racial and ethnic disparities in health care. Washington, DC: National Academy Press; 2003.
3. Fullilove MT. Comment: Abandoning "Race" As a variable in public health research—an idea whose time has come. Am J Public Health 1998;88(9):1297–1298. [PubMed: 9736864]
4. Kaufman JS, Cooper RS. Commentary: Considerations for use of racial/ethnic classification in etiologic research. Am J Epidemiol 2001;154(4):291–298. [PubMed: 11495850]
5. Krieger N. Counting accountably: Implications of the new approaches to classifying race/ethnicity in the 2000 census. Am J Public Health 2000;90(11):1687–1689. [PubMed: 11076230]
6. Jones CP, LaVeist TA, Lillie-Blanton M. "Race" In the epidemiologic literature: An examination of the American Journal of Epidemiology, 1921–1990. Am J Epidemiol 1991;134(10):1079–1084. [PubMed: 1746518]
7. Jones CP. Race, racism, and the practice of epidemiology. Am J Epidemiol 2001;154(4):299–304. [PubMed: 11495851]
8. LaVeist TA. Beyond dummy variables and sample selection: What health services researchers ought to know about race as a variable. Health Serv Res 1994;29(1):1–16. [PubMed: 8163376]
9. Osborne NG, Feit MD. The use of race in medical research. JAMA 1992;267(2):275–279. [PubMed: 1727527]
10. Sankar P, Cho MK, Condit CM, et al. Genetic research and health disparities. JAMA 2004;291(24):2985–2989. [PubMed: 15213210]
11. Stolley PD. Race in epidemiology. Int J Health Serv 1999;29(4):905–909. [PubMed: 10615582]

12. Williams DR. Race and health: Basic questions, emerging directions. AEP 1997;7(5):322–333. [PubMed: 9250627]
13. Witzig R. The medicalization of race: Scientific legitimization of a flawed social construct. Ann Intern Med 1996;125(8):675–679. [PubMed: 8849153]
14. National Institute of Health. NIH guidelines on the inclusion of women and minorities as subjects in clinical research. [Accessed 01/01/05]. http://grants.nih.gov/grants/guide/notice-files/not94-100.html
15. NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research—Amended. Oct2001 [Accessed 01/01/05]. http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm
16. Kaplan JB, Bennett T. Use of race and ethnicity in biomedical publication. JAMA 2003;289(20): 2709–2716. [PubMed: 12771118]
17. Williams DR. The concept of race in health services research: 1966 to 1990. Health Serv Res 1994;29(3):261–274. [PubMed: 8063565]
18. Glaser, BG.; Strauss, AL. The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Publishing Co; 1967. The constant comparative method of qualitative analysis; p. 101-115.
19. Hogenraad R, McKenzie D, Peladeau N. Force and influence in content analysis: The production of new social knowledge. Qual Quant 2003;37:221–238.
20. Ryan G. Measuring the typicality of text: Using multiple coders for more than just reliability and validity checks. Hum Organ 1999;58(3):313–321.
21. Patton, MQ. Qualitative evaluation and research methods. 2. Newbury Park, CA: Sage Publications; 1990.
22. Afshari R, Bhopal RS. Changing pattern of use of ‘ethnicity’ and ‘race’ in scientific literature. Int J Epidemiol 2002;31(5):1074. [PubMed: 12435788]
23. Ahdieh L, Hahn RA. Use of the terms ‘race’, ‘ethnicity’, and ‘national origins’: A review of articles in the american journal of public health, 1980–1989. Ethn Health 1996;1(1):95–98. [PubMed: 9395552]
24. Fuller KE. Health disparities: Reframing the problem. Med Sci Monit 2003;9(3):SR9–15. [PubMed: 12640358]
25. Foster MW, Sharp RR. Race, ethnicity, and genomics: Social classifications as proxies of biological heterogeneity. Genome Res 2002;12(6):844–850. [PubMed: 12045138]
26. Bowcock AM, Kidd JR, Mountain JL, et al. Drift, admixture, and selection in human evolution: A study with DNA polymorphisms. Proc Natl Acad Sci U S A 1991;88(3):839–843. [PubMed: 1992475]
27. Bowcock AM, Ruiz-Linares A, Tomfohrde J, et al. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 1994;368(6470):455–457. [PubMed: 7510853]
28. Calafell F, Shuster A, Speed WC, et al. Short tandem repeat polymorphism evolution in humans. Eur J Hum Genet 1998;6(1):38–49. [PubMed: 9781013]
29. Mountain JL, Cavalli-Sforza LL. Multilocus genotypes, a tree of individuals, and human evolutionary history. Am J Hum Genet 1997;61(3):705–718. [PubMed: 9326336]
30. Rosenberg NA, Pritchard JK, Weber JL, et al. Genetic structure of human populations. Science 2002;298(5602):2381–2385. [PubMed: 12493913]
31. Stephens JC, Schneider JA, Tanguay DA, et al. Haplotype variation and linkage disequilibrium in 313 human genes. Science 2001;293(5529):489–493. [PubMed: 11452081]
32. Judson R, Salisbury B, Schneider J, et al. How many snps does a genome-wide haplotype map require? Pharmacogenomics 2002;3(3):379–391. [PubMed: 12052145]
33. Dean M, Stephens JC, Winkler C, et al. Polymorphic admixture typing in human ethnic populations. Am J Hum Genet 1994;55(4):788–808. [PubMed: 7942857]
34. Smith MW, Lautenberger JA, Shin HD, et al. Markers for mapping by admixture linkage disequilibrium in african american and hispanic populations. Am J Hum Genet 2001;69(5):1080–1094. [PubMed: 11590548]

35. Braun L. Race, ethnicity, and health: Can genetics explain disparities? Perspect Biol Med 2002;45(2):159–174. [PubMed: 11919376]
36. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. N Engl J Med 2004;351(20):2049–2057. [PubMed: 15533851]
37. Schwartz RS. Racial profiling in medical research. N Engl J Med 2001;344(18):1392–1393. [PubMed: 11333999]
38. Wood AJ. Racial differences in the response to drugs—pointers to genetic differences. N Engl J Med 2001;344(18):1394–1396. [PubMed: 11336055]
39. van den Oord EJ, Rowe DC. Racial differences in birth health risk: A quantitative genetic approach. Demography 2000;37(3):285–298. [PubMed: 10953804]
40. Van Den Oord EJ, Rowe DC. A step in another direction: Looking for maternal genetic and environmental effects on racial differences in birth weight. Demography 2001;38(4):573–576. [PubMed: 11723954]
41. Frank R. The misuse of biology in demographic research on racial/ethnic differences: A reply to van den oord and rowe. Demography 2001;38(4):563–567. [PubMed: 11723952]
42. Zuberi T. One step back in understanding racial differences in birth weight. Demography 2001;38(4):569–571. [PubMed: 11723953]
43. Duster T. Medicine. Race and reification in science. Science 2005;307(5712):1050–1051. [PubMed: 15718453]
44. Anand SS. Using ethnicity as a classification variable in health research: Perpetuating the myth of biological determinism, serving socio-political agendas, or making valuable contributions to medical sciences? Ethn Health 1999;4(4):241–244. [PubMed: 10705561]
45. Chaturvedi N. Ethnicity as an epidemiological determinant—crudely racist or crucially important? Int J Epidemiol 2001;30(5):925–927. [PubMed: 11689494]
46. Senior PA, Bhopal R. Ethnicity as a variable in epidemiological research. Br Med J 1994;309(6950):327–330. [PubMed: 8086873]
47. Lillie-Blanton M, Laveist T. Race/ethnicity, the social environment, and health. Soc Sci Med 1996;43(1):83–91. [PubMed: 8816013]
48. Corbie-Smith G, St George DM, Moody-Ayers S, et al. Adequacy of reporting race/ethnicity in clinical trials in areas of health disparities. J Clin Epidemiol 2003;56(5):416–420. [PubMed: 12812814]

Table 1

Principal investigator (PI) characteristic (n=33)

	n	%
Minority Focused Research		
Yes	22	67
No	11	33
Academic Rank		
Assistant professor	7	21
Associate professor	8	24
Full professor	11	33
Other (Data missing for 1 PI)	6	18
Number of Published Articles		
Range	0–154	
Mean	33.3	
PI Race		
White	22	67
Black	11	33
Other	0	0
PI Sex		
Male	12	36
Female	21	64
Professional Degree		
MD	12	36
Other doctoral degree	19	58
Other	2	6
HBCU	6	18
Non-HBCU	27	82

Table 2

Themes and Illustrative Quotes

Measurement and Race
Self-Report
” …felt that was important to allow the respondents to self-report. So [the research assistant] had a check list for them. And I think that was a very good idea because, in fact there were a couple of cases where the self-report was different from the information that we already had from the other two sources.”
Generational Measures
” … what we do is we ask the individual what they consider themselves, and we give them a simplified version of the race/ethnicity list … . Then we ask them what their parents and their grandparents considered themselves, OK? Then we ask where all 3 generations were born. Now, there’s also an unknown category. Some people don’t know who their grandparents were.”
Racial/Ethnic Identity
” … we use a number of measures that get a little bit more in racial or ethnic identity, that look at how many of your friends are from the same racial group, or different racial groups, how much time do you spend with them, what do you do with them. It’s a series of like 8 questions that really gets more at racial or ethnic identity versus just race itself.”
Biological Construct
Genetic Variants
” And there turned out to be some biological differences between Hispanics and African Americans that actually do affect folic metabolism. And there’s a gene that needs you to eat more folic acid that African Americans almost never have and around 25% of Hispanics actually have.”
Differences in Response to Treatment
” [Another investigator on a panel] had a genomic project there, did the old “well, they’re all 95% the same and so on and so forth” and “oh, it doesn’t make any difference and there’s more variation within racial than there is between” and so on and so on. And, I’m going “yeah, but when we can predict a better drug response on the basis of genes and alleles that African Americans carry that whites don’t—what’s the problem here? OK, so, yes, the genetic, uh, make up between races is, is indeed more varied but the very fact that there are marker genes …” *(sic)*
Social Construct
Differences in Risk Due to Social Factors
” I think race is primarily a social construct … when we do multivariate analyses controlling for education, income employment, some of the major things I think need to be included, sometimes race is still there. So there are other parts of this social construct that is sometimes difficult to totally control for.”
Race Working through Discrimination
” But I really think of it as kind of a proxy for the experience of discrimination, the experience of oppression in this country, OK? So, kind of along the lines of how Camera Jones conceptualizes race. I think of it as a proxy for the experience of oppression in this country … I started looking at, like, feminist theory and using a feminist framework to try to locate black women in this system of multiple oppression based on race, class and gender and, and, and conceptualize it that way, try to look at the multiple oppressions that influence black women’s decision making abilities, or really takes away black women’s decision making power … . It was an issue of there’s something about being black in this country that predisposes people to an experience of oppression. There’s something about being a woman in this country that predisposes people to this experience. And, and that all of these experiences form, these intersecting oppressions create a unique experience for black women. And I wanted to understand how it was that that unique experience contributes to their increased risk for [disease].”
” I think that the influence of race on health is almost entirely through racism and that when I am looking at race, what I’m looking at is racism because the one thing that all black people have in common is some, is that they have experienced some element of racism. Whether they believe it or not, they don’t have to believe it. And, in fact, their experiences are likely different. I mean, my experience is not the same as somebody in, you know, inner-city Baltimore, but nonetheless, when I stand on a cab, stand on the street in New York City with my hand out, whether I’m wearing a Rolex and a fur coat, I still can’t get a cab! I mean, and neither can Ed Bradley.”
Social and Biological
Biologic with Modification by Social Factors
” Is it really hypertension, or are there other reasons? Genetic, dietary, cultural, activity, sedentary lifestyle, that kind of stuff. And so, you factor in nutritional and genetic and maybe even environmental factors and then also educational factors in terms of how that interacts … . And so, the socioeconomics of it amplify what may be a genetic predisposition.”
” …the amount of differences between the different races is a very, very, very small percentage of our genetic material. And so, having said that, I think that there are things that influence race over the years in terms of diet and cultures and behaviors and things like that. So, when I think of race, I don’t necessarily think of it as necessarily a genetic thing. I guess you would call it “epi-genetics” or something.”
Social Construct with Biologic Elements

” I think the way that I look at race is largely a social construct as opposed to more of a biological one, although I think there is a biological element. But largely a social construct but such a fundamental one in the United States, and most places in the world, but particularly in the United States, that it really is critical.”

” I don’t conceive of race so much as genetics, although I know that the genetics can vary somewhat. I think that my bias is the genetics probably have a role to play but it’s not as strong a role as other factors.”

Equal Influence of Social and Biologic Factors

” …the particular area I work in, I mean race is an extremely interesting variable … there’s the suggestion that, perhaps overall the quality of life is better [for African Americans with this disease] how much is this due to physiologic differences, how much is perhaps related to different ways of interpreting symptoms. There’s some very interesting literature, as I’m sure you know, on some of those questions.”

” Probably the case is that there’s an interaction … we see ethnic differences in pain perception, let’s say; I’m sure there’s both a social, cultural and a biological factor. You know, I don’t think it’s either/or.”

Challenges and Opportunities

Disaggregating Race from Other Factors

” …I don’t think anybody really knows what race is and what it represents.”

” the question that comes up is that is a difference that’s been seen, really attributable to a racial issue as it pertains to genetics or whether it’s a socioeconomic or an access to care issue. And we’ve been pretty careful about trying to dissect those, although that’s been just incredibly hard to do.”

” And, so racial disparity in this, in this illness is a big focus of our research. So, in terms of looking at race to me, you know, in some respects it’s a useless variable in that it encompasses potentially so many different things. But it at least gives you a starting point to try to separate out what those are.”

Heterogeneity of Racial Groups

” But I think that people think of minorities … as a homogenous group more or less and don’t account for the differences … so, I mean I think that, you know, based on class there’s a lot of variety. Based on culture there’s a lot of variety among people of color. Even if you just say “among black people” and you need to take those kinds of issues into consideration. And, you know, in terms of ethnicity.”

” But what we look at is the heterogeneity within black communities. My favorite variable is ethnic identity, so we’ve spent a good deal of time developing measures to say how does ethnic identity mediate culture. So, those who are high on pro-black or high on assimilation respond differently than some of our analyses, to our interventions or have different baseline characteristics. So, looking for the heterogeneity within, looking at it as a cultural and not a racial group … I think we have to really look for a diversity within groups and look for the socioeconomic and psychological variables that distinguish folks.”

Potential for Misuse/Misinterpretation

” I think, despite the absurdity of race as a biological variable, as a marker for culture and ethnicity and a possible political tool, and the fact is that these racial, cultural, ethnic groups often do have different health risks, I think we can’t ignore them, and we have to play it very carefully not to get into blaming or genetic inferiority or anything like that … You have to be real careful and real aware of some of the controversy and the pitfalls of constructing these categories.”

Table 3

Description of manuscripts (n=50)

# Manuscripts per Author, Mean (Range)	1.72 (0–2)
Guidance on Use of Race in Instructions for Authors	0[1]
Race Mentioned at Any Point in Manuscript	42 (84%)
Title	21 (42%)
Introduction	30 (60%)
Methods	30 (60%)
Tables or results	33 (66%)
Discussion	29 (58%)
Race defined or explained in manuscript	9 (18%)
Conceptualization of race	13 (26%)
Social construct	7 (14%)
Biological construct	4 (8%)
Combined	2 (4%)
Could not be determined	37 (74%)
Concordance between interview and manuscript*	9/13=full match 3/13=partial match 1/13=no match[2]

* Of those manuscripts where conceptualization of race could be determined (n=13);

[1] 10% (4 of 42) referred to the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, which asks for description of the sample, suggests avoiding the term “race” and asks for investigators to “specify what the descriptors mean”;

[2] Full match is defined as having the same conceptualization of race in the interview and in both of the manuscripts, partial match as having the same conceptualization of race in the interview and 1 out of the 2 manuscripts, and no match as having the different conceptualization of race in the interview and the manuscripts.